Inverse Problems **15** (1999) 155–173. Printed in the UK PII: S0266-5611(99)95517-X

Inverse problems related to crystallization of polymers

Martin Burger, Vincenzo Capasso† and Heinz W Engl

Industrial Mathematics Institute, Johannes Kepler Universität Linz, Altenbergerstrasse 69, A-4040 Linz, Austria

Received 2 July 1998

Abstract. This paper is devoted to identification problems in polymer crystallization processes. As a first step the identification of nucleation and growth rates from boundary measurements will be discussed in the case of one-dimensional crystallization. This leads to an identification problem in a coupled system of a hyperbolic and a parabolic equation.

An algorithm for the solution of this inverse problem will be developed using Landweber iteration as a regularization method. The behaviour of the algorithm will be studied both numerically and mathematically on the basis of the recently developed convergence theory for iterative regularization methods for nonlinear inverse problems.

1. Introduction

Mathematical modelling has been very successfully applied to predicting (and eventually optimizing) the final morphology of crystallized polymers under various operating conditions (cf e.g. [2, 11, 12, 20, 22]). The quality of the predictions made by mathematical models depends crucially on the use of accurate values for the various physical parameters that appear in the model. Many of these parameters are not accessible to direct measurements and have thus to be determined by 'indirect measurements', i.e. by parameter identification techniques. As a general reference for parameter identification we refer to [3]. Since parameter identification is an ill-posed inverse problem, major numerical difficulties arise due to the inherent instability of such problems (see [13] for general background material on inverse problems and regularization methods for their solution). This issue is especially important due to the fact that the available measurements are rather noisy.

Our aim in this paper is to construct a stable algorithm for identifying temperature-dependent growth and nucleation rates in a recently developed model for one-dimensional crystallization (cf [6]). This algorithm is based on Landweber iteration; for this method, applied to nonlinear inverse problems, a comprehensive convergence theory is available (cf [10, 18]). We apply this theory to our model in a formal way, i.e. we do not verify certain theoretical conditions like e.g. the existence of solutions for the direct problem in certain Hilbert spaces, since this would be technically rather involved and would obscure the basic ideas. Of course, for a complete justification of our method, this will eventually have to be done in future work.

† On leave from (permanent address): Dipartimento di Matematica, Università di Milano, Via Saldini 50, I-20133 Milano, Italy.

0266-5611/99/010155+19$19.50

The model for one-dimensional crystallization of polymers developed in [6] leads to initial-boundary value problems of the form

$$\frac{\partial T}{\partial t} = D\frac{\partial^2 T}{\partial x^2} + L\frac{\partial \xi}{\partial t} \tag{1.1}$$

$$\frac{\partial}{\partial t}\left(\frac{1}{\tilde{G}(T)(1-\xi)}\frac{\partial \xi}{\partial t}\right) = \frac{\partial}{\partial x}\left(\frac{\tilde{G}(T)}{1-\xi}\frac{\partial \xi}{\partial x}\right) + 2\frac{\partial}{\partial t}(\tilde{N}(T)) \tag{1.2}$$

with initial conditions

$$T(x,0) = T^0(x) \tag{1.3}$$

$$\xi(x,0) = 0 \tag{1.4}$$

$$\frac{\partial \xi}{\partial t}(x,0) = 0 \tag{1.5}$$

and boundary conditions

$$\frac{\partial T}{\partial n}(x,t) = \alpha(T(x,t) - T^1(x,t)) \qquad \text{for } x \in \partial\Omega \tag{1.6}$$

$$\frac{\partial \xi}{\partial t}(x,t) + \tilde{G}(T)\frac{\partial \xi}{\partial n}(x,t) = 0 \qquad \text{for } x \in \partial\Omega \tag{1.7}$$

where T^1 in (1.6) represents an exterior temperature in the cooling process at the boundary.

In these equations, T denotes the temperature, ξ the degree of crystallinity. The parameter arising in the heat transfer problem, i.e. the diffusion coefficient D, the latent heat L and the heat transition coefficient α, can be determined experimentally. The function $\tilde{N}$ represents an equivalent nucleation rate per unit of length (cf [6]) and $\tilde{G}$ the radial growth rate of a nucleus. Note that $\tilde{G}$ and $\tilde{N}$ depend on temperature rather strongly and cannot (at least not simultaneously) be measured as functions of temperature. Some materials admit experimental determination of the growth rate (cf [22]), in these cases the identification of $\tilde{N}$ for known $\tilde{G}$ is of special interest.

The measurable quantities in usual experiments are the temperature T on the boundary of the crystallization domain Ω (or at least on a part of the boundary $\Gamma \subset \partial\Omega$) as well as the degree of crystallinity $\xi(x, t_*)$ at the end of the process. The temperature at the end of the experiment is not measurable, because the structure is frozen in by a final quench below the melting point.

Furthermore, data about the final morphology, i.e. about the final distribution of nuclei $\Lambda(x, t_*)$, are sometimes available, this quantity being linked to the nucleation rate via

$$\Lambda(x,t_*) = \int_0^{t_*} (1-\xi(x,t))\frac{\partial \tilde{N}(T)}{\partial t}\,\mathrm{d}t. \tag{1.8}$$

The use of such data is different from the use of boundary data and will be discussed in section 2.4.

A major difference to parameter estimation problems in parabolic and hyperbolic differential equations with overspecified boundary data (cf e.g. [7, 15, 19, 21]) is the strong coupling of the parabolic and the hyperbolic equation. The fact that the parameters $\tilde{N}$ and $\tilde{G}$ are functions of the temperature, but arise in the hyperbolic equation, is another unusual feature of this identification problem.

1.1. Simplification and scaling

System (1.1), (1.2) may be transformed into a system of PDEs of first order in time given by

$$\frac{\partial T}{\partial t} = D\frac{\partial^2 T}{\partial x^2} + L\frac{\partial \xi}{\partial t} \tag{1.9}$$

$$\frac{\partial \xi}{\partial t} = (1-\xi)\tilde{G}(T)w \tag{1.10}$$

$$\frac{\partial w}{\partial t} = \frac{\partial}{\partial x}\left(\frac{\tilde{G}(T)}{1-\xi}\frac{\partial \xi}{\partial x}\right) + 2\frac{\partial \tilde{N}(T)}{\partial t}. \tag{1.11}$$

In practical applications the crystallization is performed under symmetric conditions, so that it is reasonable to assume symmetric boundary and initial conditions in the one-dimensional domain $\Omega = (a, b)$. Also by symmetry we may restrict the problem to $(a, \frac{a+b}{2})$ and use the conditions

$$\frac{\partial T}{\partial x}\left(\frac{a+b}{2}, \cdot\right) = 0$$

$$\frac{\partial \xi}{\partial x}\left(\frac{a+b}{2}, \cdot\right) = 0$$

at the new right boundary.

By using the quantity $v = -\ln(1-\xi)$, mapping $[a, \frac{a+b}{2}]$ to $[0, 1]$ and further scaling the system may be transformed into the form

$$\begin{aligned} u_t &= Du_{xx} + L\mathrm{e}^{-v}v_t \\ v_t &= a(u)w \\ w_t &= (a(u)v_x)_x + b(u)_t \end{aligned} \tag{1.12}$$

$$\begin{aligned} u|_{t=0} &= u^0 \\ v|_{t=0} &= 0 \\ w|_{t=0} &= 0 \end{aligned} \tag{1.13}$$

$$\begin{aligned} u_x(0,t) &= -\alpha(u(0,t) - u^1(0,t)) \\ w(0,t) - v_x(0,t) &= 0 \end{aligned} \tag{1.14}$$

$$\begin{aligned} u_x(1,t) &= 0 \\ v_x(1,t) &= 0. \end{aligned} \tag{1.15}$$

The quantities $a(u)$ and $b(u)$ in (1.12) are the equivalents to $\tilde{G}(u)$ and $\tilde{N}(u)$, transformed to the new reference variable u. The function u is a scaled temperature, i.e. $u = 1$ is above the equilibrium melting point and $u = 0$ below the glass transition temperature, hence the temperature range of interest $[u_1, u_2]$ is a subset of $[0, 1]$. In practice, the temperature at the beginning of the experiment is always above the melting point, i.e. u_2 may be chosen as the melting point. The fact that no nucleation occurs above this temperature implies that $b(u) = 0$ for $u \geqslant u_2$, so we will restrict our attention to functions $b \in H^1([u_1, u_2])$ satisfying $b(u_2) = 0$.

In the following we will assume that the temperature data are given at the point $x = 0$ by $u(0, t) = u_B(t)$ in the time interval $I = (0, t_*)$. From the available observation of ξ at $t = t_*$ we can compute the value v_* for $v|_{t=t_*}$, which we therefore assume to be given.

By the remarks made above, identification of the nucleation rate $\tilde{N} = \tilde{N}(T)$ in (1.1)–(1.7) is equivalent to identifying $b = b(u)$ in (1.12)–(1.15). Thus, this identification problem can formally be written as the nonlinear operator equation

$$F(b) = (u_B^\delta, v_*^\delta), \tag{1.16}$$

where u_B^δ and v_*^δ denote noisy measurements of u_B and v_* with noise level δ. The operator F, the 'parameter-to-output-map', maps an admissible value of the parameter b to the values u_B and v_* obtained from solving (1.12)–(1.15) with this parameter. In the next section we treat the nonlinear operator equation (1.16) in a more formal way and develop a stable algorithm for its solution.

2. Identification of nucleation rates

We assume that for 'admissible' parameter $b \in \mathcal{D}(F)$ the direct problem (1.12)–(1.15) admits a unique strong solution

$$(u, v, w) \in X := H^{1,2}(Q) \times H^{2,2}(Q) \times H^{1,0}(Q)$$

which depends continuously on b in the H^1-norm, where the time-space domain Q is defined by

$$Q := \Omega \times I = (0, 1) \times (0, t_*).$$

This well-posedness requirement can also be thought of as an implicit definition of $\mathcal{D}(F)$. Finding conditions on b that guarantee the well-posedness of the nonlinear system (1.12)–(1.15) is in principle standard, but technically quite involved. Under this well-posedness assumption, the operator used in (1.16) can be formally defined as a continuous operator

$$\begin{aligned} &F : \mathcal{D}(F) \subset V := \{b | b \in H^1([u_1, u_2]), b(u_2) = 0\} \to L^2(I) \times L^2(\Omega) \\ &b \mapsto (u|_{x=0}, v|_{t=t_*}). \end{aligned} \tag{2.1}$$

The inverse problem of solving (1.16) is most likely ill-posed, so that regularization has to be used. One classical possibility would be Tikhonov regularization, i.e. taking as regularized solution (global) minimizers of the functional

$$b \mapsto \|F(b) - (u_B^\delta, v_*^\delta))\|^2_{L^2(I)\times L^2(\Omega)} + \alpha \|b\|^2_{H^1([u_1,u_2])} \tag{2.2}$$

over $\mathcal{D}(F)$, where $\alpha > 0$ is a regularization parameter. The convergence theory for this method (for nonlinear problems) has been developed in [14] (cf also [13]). A major disadvantage, however, is that since the functional in (2.2) is in general not convex and might have many local minima, minimizing (2.2) is not an easy task. Therefore, iterative regularization methods are an attractive alternative especially for nonlinear ill-posed problems; the regularization effect comes from stopping the iteration at an appropriately defined stopping index which depends also on the noise level δ ('stopping rule'). The convergence theory of iterative methods for solving nonlinear ill-posed problems is still a developing field of research (see [10, 13, 16] and the references quoted there).

Here we use one of the simplest iterative regularization methods, namely Landweber iteration, which nevertheless turned out to be quite effective for certain (especially severely) ill-posed nonlinear problems (cf [17]). In later work we will also investigate the efficiency of using Newton-like methods like the iteratively regularized Gauss–Newton method (cf [5, 10]), which is of course faster than Landweber for data without noise, but also amplifies data noise faster.

For our problem (1.16), Landweber iteration is defined by

$$b^{k+1} = b^k - \omega F'(b^k)^*(F(b^k) - (u_B^\delta, v_*^\delta)). \tag{2.3}$$

For computing the iterates, the adjoint $F'(b^k)^*$ is needed. As we shall see later, the adjoint also plays a crucial role in the analysis of the convergence rate. Hence, it is necessary to be able to compute these adjoints efficiently. Note that the adjoint depends on the Hilbert spaces between which F acts, so that also the choice of these spaces is important. In the following we derive a method of computing the adjoint of F' by solving a system of linear partial differential equations. We first split F into $F = \Psi \circ \Phi$, where

$$\begin{aligned} &\Phi : \mathcal{D}(F) \subset H^1([u_1, u_2]) \to X \\ &b \mapsto (u, v, w) \end{aligned}$$

is the parameter-to-solution map and

$$\Psi : \Phi(\mathcal{D}(F)) \subset X \to L^2(I) \times L^2(\Omega)$$
$$(u, v, w) \mapsto (u|_{x=0}, v|_{t=t_*})$$

is the trace operator that maps the solution onto u at $x = 0$ and v at $t = t_*$. As Ψ is a linear and continuous operator, the Fréchet derivative of F is $F' = \Psi \circ \Phi'$. We assume that Φ' exists as a Fréchet derivative (verifying this under appropriate smoothness conditions on a and b would certainly be possible, but quite technical) and proceed in a formal way.

By straightforward linearization we conclude that the derivative $(U, V, W) := \Phi'(b)h$ satisfies the system

$$\begin{aligned} U_t &= DU_{xx} + L\mathrm{e}^{-v}(v_t V + V_t) \\ V_t &= a(u)W + a'(u)U w \\ W_t &= (a(u)V_x)_x + (a'(u)U v_x)_x + (b'(u)U)_t + h(u)_t \end{aligned} \tag{2.4}$$

with initial and boundary conditions given by

$$\left.\begin{aligned} \alpha U + U_x &= 0 \\ W - V_x &= 0 \end{aligned}\right\} \quad \text{for } x = 0 \tag{2.5}$$

$$\left.\begin{aligned} U_x &= 0 \\ V_x &= 0 \end{aligned}\right\} \quad \text{for } x = 1 \tag{2.6}$$

$$\begin{aligned} U|_{t=0} &= 0 \\ V|_{t=0} &= 0 \\ W|_{t=0} &= 0. \end{aligned} \tag{2.7}$$

System (2.4)–(2.7) is linear in (U, V, W), but depends on the current value of b and hence on the solution $(u, v, w) = \Phi(b)$ of the nonlinear system (1.12)–(1.15). We again assume that the solution of (2.4)–(2.7) exists and is unique. From this system, we can compute the derivative of F as

$$F'(b)h = \Psi \circ \Phi'(b)h. \tag{2.8}$$

2.1. The adjoint problem

For the computation of the adjoint we split $F'(b)$ as

$$F'(b)h = S \circ R \circ Jh \tag{2.9}$$

where J is the embedding operator from V onto $L^2([u_1, u_2])$, R the operator

$$R : \mathcal{D}(F) \subset L^2([u_1, u_2]) \to L^2(\Omega \times I)$$
$$h(u) \mapsto \overline{h}(x, t) := h(u(x, t))$$

and S is the operator defined by

$$S : L^2(\Omega \times I) \times L^2(I) \to L^2(I) \times L^2(\Omega)$$
$$\overline{h} \mapsto (U|_{x=0}, V|_{t=t^*})$$

by which we mean the trace of the solution of (2.4)–(2.7) with $h(u)_t$ replaced by $\frac{\partial \overline{h}}{\partial t}(x, t)$ in the last equation and b being the argument in $F'(b)$.

Because of (2.9),

$$F'(b)^* = J^* \circ R^* \circ S^*. \tag{2.10}$$

We compute these three adjoints step by step, starting with S^*. This will involve the following problem:

$$\begin{aligned}\phi_t &= -D\phi_{xx} - a'(u)\psi w - a(u)b'(u)\psi + a'(u)v_x\theta_x\\ \psi_t &= L\mathrm{e}^{-v}\phi_t - (a(u)\theta_x)_x\\ \theta_t &= -a(u)\psi\end{aligned} \tag{2.11}$$

$$\begin{aligned}\phi|_{t=t_*} &= 0\\ \psi|_{t=t_*} &= r\\ \theta|_{t=t_*} &= 0\end{aligned} \tag{2.12}$$

$$\left.\begin{aligned}D\alpha\phi + D\phi_x - a'(u)v_x\theta &= q\\ \psi - \theta_x &= 0\end{aligned}\right\} \qquad \text{at } x = 0 \tag{2.13}$$

$$\left.\begin{aligned}\phi_x &= 0\\ \theta_x &= 0\end{aligned}\right\} \qquad \text{at } x = 1 \tag{2.14}$$

considered as an equation for (ϕ, ψ, θ). Note that via u and v, the coefficients in this system depend on b. Again we assume (proceeding still in a merely formal way) unique solvability of (2.11)–(2.14).

Now let (ϕ, ψ, θ) be the solution of (2.11)–(2.14). In (2.3), (q, r) will be the residual

$$(q, r) = F(b) - (u_B^\delta, v_*^\delta) \tag{2.15}$$

but the argument holds for any $(q, r) \in L^2(I) \times L^2(\Omega)$.

Integration by parts yields

$$\begin{aligned}\int_0^{t_*}\int_0^1 \overline{h}a(u)\psi \,\mathrm{d}x\,\mathrm{d}t &= -\int_0^{t_*}\int_0^1 \overline{h}\theta_t \,\mathrm{d}x\,\mathrm{d}t\\ &= \int_0^{t_*}\int_0^1 \overline{h}_t\theta \,\mathrm{d}x\,\mathrm{d}t - \int_0^1 (\overline{h}\theta)|_{t=0}^{t=t_*}\,\mathrm{d}x.\end{aligned}$$

The assumptions $\overline{h}(x, 0) = h(u(x, 0)) = h(u^0(x)) = 0$ and $\theta|_{t=t_*} = 0$ cause the second integral to vanish. Using system (2.4) we conclude that

$$\int_0^{t_*}\int_0^1 \overline{h}_t\theta \mathrm{d}x\,\mathrm{d}t = \int_0^{t_*}\int_0^1 (W_t - (a(u)V_x)_x - (a'(u)Uv_x)_x - (b'(u)U)_t)\theta \mathrm{d}x\,\mathrm{d}t$$

$$0 = \int_0^{t_*}\int_0^1 (U_t - DU_{xx} - L\mathrm{e}^{-v}(v_tV + V_t))\phi \mathrm{d}x\,\mathrm{d}t$$

$$0 = \int_0^{t_*}\int_0^1 (V_t - a(u)W - a'(u)Uw)\psi \mathrm{d}x\,\mathrm{d}t.$$

By adding these equations and integrating by parts we finally deduce

$$\begin{aligned}\int_0^{t_*}\int_0^1 \overline{h}(x,t)a(u(x,t))\psi(x,t)\mathrm{d}x\,\mathrm{d}t &= \int_0^{t_*}\int_0^1 \overline{h}_t(x,t)\theta(x,t)\mathrm{d}x\,\mathrm{d}t\\ &= \int_0^{t_*} U(0,t)q(t)\,\mathrm{d}t + \int_0^1 V(x,t_*)r(x)\,\mathrm{d}x\end{aligned}$$

hence

$$S^*(q, r) = a(u)\psi \tag{2.16}$$

where ψ is defined via (2.11)–(2.14).

The adjoint of J can be computed by solving

$$-\sigma''(u) + \sigma(u) = \rho(u) \tag{2.17}$$

$$\sigma'(u_1) = 0 \qquad \sigma(u_2) = 0 \tag{2.18}$$

for $\sigma = J^*\rho$: The unique solution σ satisfies $\sigma \in V$ and, for any $b \in V$,

$$\begin{aligned}(b,\sigma)_V = (b,\sigma)_{H^1} &= \int_{u_1}^{u_2} (b'(u)\sigma'(u) + b(u)\sigma(u))\,\mathrm{d}u \\ &= \int_{u_1}^{u_2} (b(u)\rho(u))\,\mathrm{d}u \\ &= (Jb,\rho)_{L^2}.\end{aligned}$$

The computation of the adjoint of R is more involved, so that we will only consider the case of decreasing temperature, which is the only one of practical interest. If $u(x,.) \in C^1(I)$ and $u_t \leqslant c < 0$ there exists a regular 'change-of-variables' transformation $p : (t,x) \mapsto (u,x)$. Because of $|\det p'(x,t)| = -u_t(x,t)$ and $u(p^{-1}(\overline{u},x)) = \overline{u}$, the substitution rule for integrals implies in this case that

$$\begin{aligned}\int_0^{t_*}\int_0^1 (Rh)(x,t)\varphi(x,t)\mathrm{d}x\,\mathrm{d}t &= \int_0^{t_*}\int_0^1 h(u(x,t))\varphi(x,t)\mathrm{d}x\,\mathrm{d}t \\ &= -\int_{u_1}^{u_2}\int_{I(\overline{u})} h(\overline{u})\varphi(p^{-1}(\overline{u},x))\frac{1}{u_t(p^{-1}(\overline{u},x))}\mathrm{d}x\,\mathrm{d}\overline{u} \\ &= \int_{u_1}^{u_2} h(\overline{u})\left(-\int_{I(\overline{u})}\frac{\varphi(p^{-1}(\overline{u},x))}{u_t(p^{-1}(\overline{u},x))}\,\mathrm{d}x\right)\mathrm{d}\overline{u}\end{aligned}$$

where

$$I(\overline{u}) = \{x \in \Omega \mid \exists t : u(x,t) = \overline{u}\}.$$

Thus, $R^*\varphi$ is given by

$$(R^*\varphi)(\overline{u}) = -\int_{I(\overline{u})}\frac{\varphi(p^{-1}(\overline{u},x))}{u_t(p^{-1}(\overline{u},x))}\,\mathrm{d}x. \tag{2.19}$$

As a straightforward calculation shows, $R^*\varphi$ is still well-defined in $L^2([u_1,u_2])$. If u is an element of $H^{1,2}(\Omega \times I)$ satisfying $u_t \leqslant c < 0$ almost everywhere, by a density argument the form (2.19) holds for such u, too. We do not go into details, since (2.19) is not directly usable numerically anyway.

2.2. An identification algorithm

Since (2.19) seems not to be usable in a direct way, we compute $J^*R^*\varphi$ approximately by a projection method. $J^*R^*\varphi$ satisfies

$$(RJp,\varphi) = [p, J^*R^*\varphi]$$

for all $p \in H^1([u_1,u_2])$, where $(.,.)$ is the inner product in $L^2(Q)$ and $[.,.]$ the one in $H^1([u_1,u_2])$. If we choose a finite-dimensional subspace V_N of $H^1([u_1,u_2])$ with basis $\{p_1,\ldots,p_N\}$, the orthogonal projection $\tilde{\varphi}_N$ of $\tilde{\varphi} := J^*R^*\varphi$ onto V_N may be written as

$$\tilde{\varphi}_N = \sum_{i=1}^N \alpha_i p_i$$

and satisfies

$$(RJp_i,\varphi) = [p_i,\tilde{\varphi}] = [p_i,\tilde{\varphi}_N] = \sum_{j=1}^N \alpha_j [p_i,p_j]$$

so we obtain the linear system

$$P\underline{\alpha} = \underline{\beta}$$

for the coefficients $\underline{\alpha} = (\alpha_1, \ldots, \alpha_N)$, where

$$P = ([p_i, p_j])_{i,j=1,\ldots,N}$$
$$\underline{\beta} = ((RJp_i, \varphi))_{i=1,\ldots,N}.$$

The integrals (Rp_i, φ) can be computed easily, although even a small support of the functions p_i does not restrict the domain of integration in general. If the functions p_i are the usual linear splines

$$p_i(s) = \begin{cases} \frac{s-s_{i-1}}{s_i-s_{i-1}} & \text{if } s_{i-1} \leqslant s \leqslant s_i \\ \frac{s_{i+1}-s}{s_{i+1}-s_i} & \text{if } s_i \leqslant s \leqslant s_{i+1} \\ 0 & \text{if } s \notin (s_{i-1}, s_{i+1}) \end{cases}$$

and $s_i = \frac{i}{N}$, P turns out to be a tri-diagonal matrix. Hence the system can be solved efficiently, a decomposition of P can be computed in a preprocessing step.

Now, we summarize our algorithm by putting all pieces, especially the various adjoints involved, together.

Algorithm. We start with an initial guess b^0, and as long as the stopping rule is not satisfied, we do the following iteration.

- Compute $F(b^k)$ by solving the nonlinear initial-boundary value problem (1.12)–(1.15) and evaluating (u_B, v_*).
- Compute the residual (q^δ, r^δ) by (2.15).
- Compute the adjoint $S^*(q, r) = a(u)\psi$ by solving the linear initial-boundary value problem (2.11)–(2.14).
- Compute $F'(b^k)^*(q, r) = J^*R^*a(u)\psi$ by the projection method just described.
- Compute the new iterate $b^{k+1} = b^k - \omega F'(b^k)^*(q^\delta, r^\delta)$ with appropriately chosen ω (cf section 2.3).

As the stopping rule, we use the discrepancy principle (cf section 2.3.1), i.e. the stopping index k_* is the first index with

$$\|F(b^k) - (u_B^\delta, v_*^\delta)\| \leqslant \tau\delta$$

where δ is an L^2-bound for the noise in our observations.

We use some fixed $\tau > 2$ in the discrepancy principle, although τ should rather be chosen according to (2.26) as explained in detail in the following section. However, the information needed to compute τ according to this theory is not easily available.

2.3. Convergence analysis

2.3.1. General theory. We summarize the basic convergence theory for Landweber iteration for a nonlinear ill-posed operator equation

$$F(x) = y^\delta \tag{2.20}$$

between Hilbert spaces X and Y (see [10, 13, 18] for details). The right-hand side y^δ represents a perturbation of the exact data y with error level δ, i.e.

$$\|y - y^\delta\| \leqslant \delta. \tag{2.21}$$

We assume that the exact data y are attainable and the starting value is sufficiently close to the solution, i.e. there exists a solution x_* of (2.20) with $\delta = 0$ in $\mathcal{B}_\rho(x_0)$.

As in the linear case it is important that the iteration given by

$$x_{k+1}^\delta = x_k^\delta - F'(x_k^\delta)^*(F(x_k^\delta) - y^\delta) \tag{2.22}$$

is properly scaled, i.e. either

$$\|F'(x)\| \leqslant 1 \qquad \forall x \in \mathcal{B}_\rho(x_0) \tag{2.23}$$

or an appropriate damping factor ω is used.

Besides existence of a solution for exact data, the following condition is important for all convergence proofs (also for other iterative methods for solving nonlinear ill-posed problems):

$$\|F(x) - F(\overline{x}) - F'(x)(x - \overline{x})\| \leqslant \eta \|F(x) - F(\overline{x})\| \qquad \forall x, \overline{x} \in \mathcal{B}_\rho(x^0) \subset \mathcal{D}(F) \tag{2.24}$$

where η is a real number satisfying $\eta < \frac{1}{2}$. It has been shown by Hanke *et al* [18] that (2.24) implies the existence and uniqueness of a solution $x^\dagger$ of (2.20) of minimal distance to the starting point x^0. For an analysis of this condition, which somehow restricts the nonlinearity of F see [4, 18, 23].

Convergence of the Landweber iteration can be shown if the stopping index $k_* = k_*(\delta, y^\delta)$ is chosen according to the generalized discrepancy principle (cf [13, 18])

$$\|F(x_{k_*}^\delta) - y^\delta\| \leqslant \tau\delta \leqslant \|F(x_k^\delta - y^\delta)\| \qquad 0 \leqslant k < k_* \tag{2.25}$$

where τ is a positive real number satisfying

$$\tau > 2\frac{1+\eta}{1-2\eta} > 2. \tag{2.26}$$

Convergence rates can, as usual for ill-posed problems even in the linear case, only be established if 'source conditions' are satisfied. If the difference between the starting iterate and the true solution satisfies

$$x^\dagger - x^0 = (F'(x^\dagger)^* F'(x^\dagger))^\nu w_1 \tag{2.27}$$

with $0 < \nu \leqslant \frac{1}{2}$, the convergence rate

$$k_* \leqslant c_1 \left(\frac{\|w_1\|}{\delta}\right)^{\frac{2}{2\nu+1}} \tag{2.28}$$

$$\|x^\dagger - x_{k_*}^\delta\| \leqslant c_2 \|w_1\|^{\frac{1}{2\nu+1}} \delta^{\frac{2\nu}{2\nu+1}} \tag{2.29}$$

can be proved under some additional assumptions (cf [18]). The stopping index can be estimated by (2.28). In the case of $\nu = \frac{1}{2}$, (2.27) is equivalent to

$$x^\dagger - x^0 = F'(x^\dagger)^* w_0. \tag{2.30}$$

Thus, the adjoint is also needed for assessing the convergence rate. Under (2.30), the convergence rate is

$$\|x^\dagger - x_{k_*}^\delta\| = \mathcal{O}(\sqrt{\delta}) \tag{2.31}$$

and the stopping index can be estimated by

$$k_* = \mathcal{O}\left(\frac{1}{\delta}\right). \tag{2.32}$$

2.3.2. Application to our identification problem. In the special case of identifying b in (1.12)–(1.15), the source condition (2.30) can be interpreted as follows. If $u_t < 0$, i.e. the temperature is decreasing, we may write

$$(R^* S^*(q, r))(\overline{u}) = -a(u) \int_{I(\overline{u})} \frac{\psi(p^{-1}(\overline{u}, x))}{u_t(p^{-1}(\overline{u}, x))} \, \mathrm{d}x \tag{2.33}$$

where ψ is defined by (2.11)–(2.14). Together with the adjoint of the embedding operator J one obtains the source condition

$$b^\dagger - b^0 = J^* R^* S^*(\tilde{q}, \tilde{r}) \tag{2.34}$$

where the last step is to solve (2.17), (2.18) with the special right-hand side $\rho = R^* S^*(q, r)$. A standard regularity argument for the homogenous two-point boundary value problem (2.17), (2.18) implies that its solution σ is in $H^{2+\alpha}([u_1, u_2])$ if $\rho \in H^\alpha([u_1, u_2])$. From (2.33) we know at least $\rho = R^* S^*(q, r) \in L^2([u_1, u_2])$, hence

$$b^\dagger - b^0 \in \{\varphi \in H^2([u_1, u_2]) | \varphi'(u_1) = 0, \varphi(u_2) = 0\} \tag{2.35}$$

is necessary for (2.34).

Since the source condition is needed for getting a fast convergence rate for the identification problem, one should, if at all possible, try to set up the problem (and the starting iterate) in such a way that it is fulfilled. This leads to the following conclusions.

• The fact that temperature should decrease during the process (cf (2.19) and the discussion below) is heuristically obvious, because one just cannot identify a function of temperature in an experiment at constant temperature. This property is usually satisfied in experiments using *differential scanning calorimetry* (cf [11]), but not in the case of isothermal experiments, which have been used for attempts of parameter identification with different methods in the past (cf [1, 8]).

• Furthermore, formula (2.34) shows that the values of the nucleation rate at temperatures that do not occur in the experiment have to be incorporated into the initial guess b^0, because $R^* S^*(p, q)$ is always zero at these values of the temperature, which is also heuristically clear. The temperature range $[T_1, T_2]$ occurring in the experiment can be determined in practice. As the temperature of the material is always controlled by cooling at the boundary, it attains its minimum there. The maximum T_2 is always the melting point of the material, because the experiment must be started at a sample temperature above the melting point.

• The smoothness condition $b^\dagger - b^0 \in H^2([u_1, u_2])$ in (2.35) means that only smooth parts of $b^\dagger$ should be computed in the iteration procedure, rougher parts should be incorporated into the initial guess b^0 in order to obtain fast convergence. The condition at the right boundary in (2.35) can be satisfied easily, because we have restricted the domain of F to V using the *a priori* knowledge $b(u_2) = 0$. Fulfilling the condition at the left boundary

$$(b^\dagger)'(u_1) = (b^0)'(u_1) \tag{2.36}$$

is much more difficult. If u_1 represents the glass transition temperature of the material we may use the *a priori* knowledge that the nucleation rate is constant at temperatures below, i.e. $b'(u) = 0$ for $u \leqslant u_1$, so we assume $b'(u_1) = 0$ for the solution and incorporate this boundary condition into the initial guess b^0. In experiments, where not the whole temperature range between melting point and glass transition temperature may occur, so that u_1 is higher than the melting point, we may expect a fast convergence rate only if we know $b'(u_1)$, which is a strong, maybe unrealistic assumption.

2.4. Data about the morphology

The identification of rates from information about the final morphology differs from the above identification problem, because (1.8) does not only contain information at the boundary or the final time, but averaged information in the whole time-space domain. This identification problem could be handled in a similar way, we just give a short sketch of the necessary ingredients.

The parameter-to-data map is given by

$$H : \mathcal{D}(H) \subset H^1([u_1, u_2]) \to L^2(\Omega) \tag{2.37}$$

$$b \mapsto \Lambda|_{t=t_*} \tag{2.38}$$

where

$$\Lambda(x,t) = \int_0^t \mathrm{e}^{-v}(x,\tau) b(u(x,s))_s \,\mathrm{d}s.$$

The derivative $H'(b)h$ is given by

$$(H'(b)h)(x) = \int_0^{t_*} \mathrm{e}^{-v}(-Vb(u) + (b'(u)U)_t + h(u)_t)\mathrm{d}t$$

where (U, V, W) is the solution of (2.4)–(2.7).

Similar to (2.11)–(2.14) we define an adjoint problem by

$$\begin{aligned}
&\phi_t = -D\phi_{xx} - a'(u)\psi w - a(u)b'(u)\psi - a'(u)v_x\theta + \mathrm{e}^{-v}v_t b'(u)p \\
&\psi_t = L\mathrm{e}^{-v}\phi_t + (a(u)W_x)_x + \mathrm{e}^{-v}p \qquad (2.39)\\
&\theta_t = -a(u)\psi
\end{aligned}$$

$$\begin{aligned}
&\phi|_{t=t_*} = 0 \\
&\psi|_{t=t_*} = 0 \qquad (2.40)\\
&\theta|_{t=t_*} = 0
\end{aligned}$$

$$\left.\begin{aligned} &D\alpha\phi - D\phi_x - a'(u)v_x\theta = b'(1-\mathrm{e}^{-v})p \\ &\theta_x + \psi = 0 \end{aligned}\right\} \quad \text{at } x = 0 \tag{2.41}$$

$$\left.\begin{aligned} &\phi_x = 0 \\ &\theta_x = 0 \end{aligned}\right\} \quad \text{at } x = 1 \tag{2.42}$$

where p represents the residual

$$p = H(b) - \Lambda^{\delta}.$$

Integration by parts yields

$$\int_0^1 (H'(b)h)p\,\mathrm{d}x = \int_0^{t_*}\int_0^1 h(u)(-\mathrm{e}^{-v}p)_t \mathrm{d}x\,\mathrm{d}t + \int_0^1 (h(u)\mathrm{e}^{-v}p)|_{t=t_*}\,\mathrm{d}x. \tag{2.43}$$

By splitting H' analogously as in section 2.1 it is possible to compute the adjoint $H'(b)p$ by first solving the linear initial-boundary value problem (2.39)–(2.42) and then computing an adjoint similary to J^*R^*. The main difference is that instead of R the operator

$$\begin{aligned}
&\tilde{R} : \mathcal{D}(H) \subset L^2([u_1, u_2]) \to L^2(\Omega \times I) \times L^2(\Omega) \\
&h(u) \mapsto (\overline{h}(x,t), h_*(x)) = h(u(x,t)), h(u(x,t_*))
\end{aligned}$$

has to be used because of the second integral at the right-hand side of (2.43).

The data about the morphology can be used together with measurements of the boundary temperature and the final degree of crystallinity: in this case the parameter-to-data map consists of three components, i.e. b is mapped to (u_B, ξ_*, Λ_*). Because of the different structures of boundary or final data and the data Λ_* about the morphology it is not possible to compute the adjoint of the parameter-to-data map needed for the Landweber iteration by solving one initial-value problem. Systems (2.11)–(2.14) and (2.39)–(2.42) must be solved separately to obtain the adjoints of the components, the adjoint of the parameter-to-data map is then just their sum. Although we did not perform numerical tests with this set-up of the problem, it seems obvious that additional information about the morphology will improve the quality of the solution.

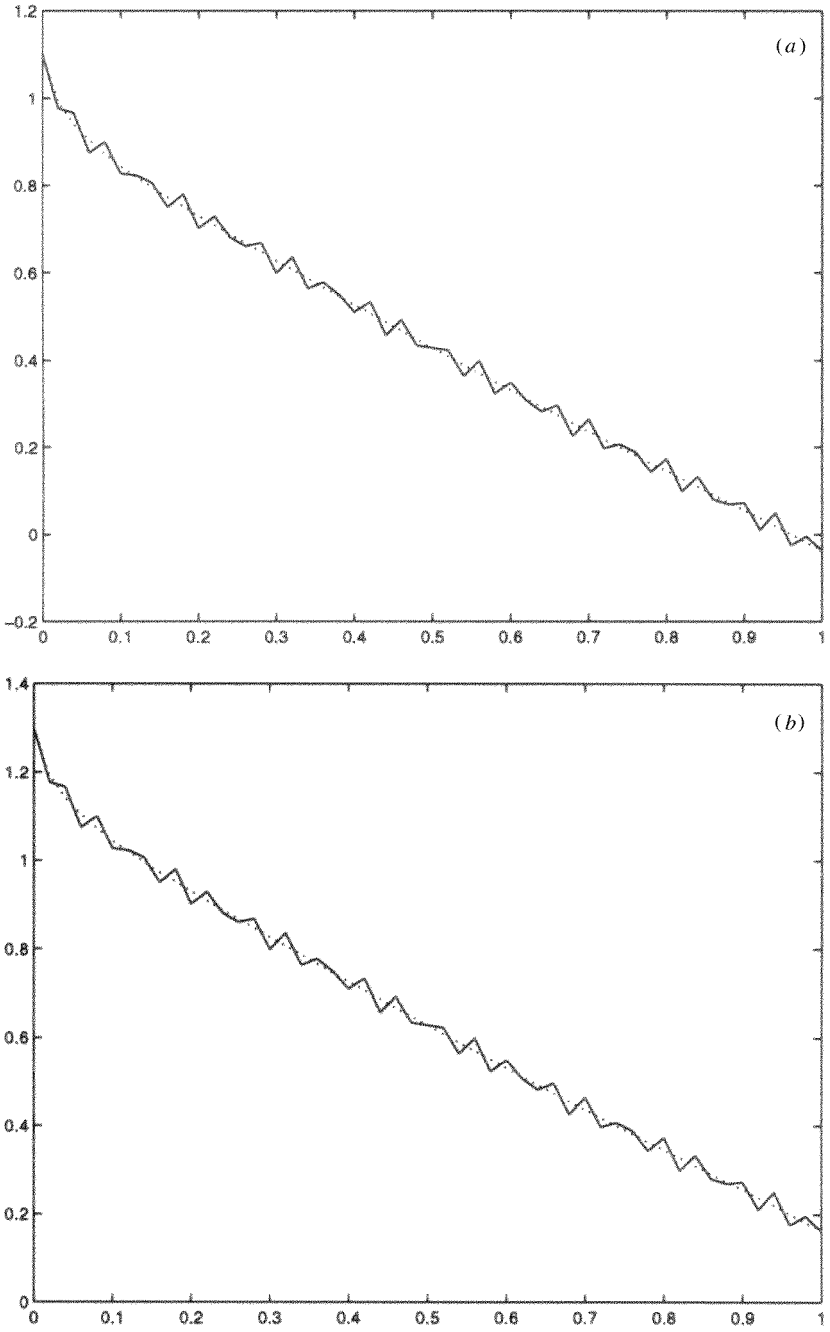

Figure 1. Exact and perturbed data for the boundary temperature, $\delta = 0.025$. (*a*) corresponds to an experiment with temperature values in the whole range $[0, 1]$; (*b*) to an experiment with temperature not decreasing below 0.2.

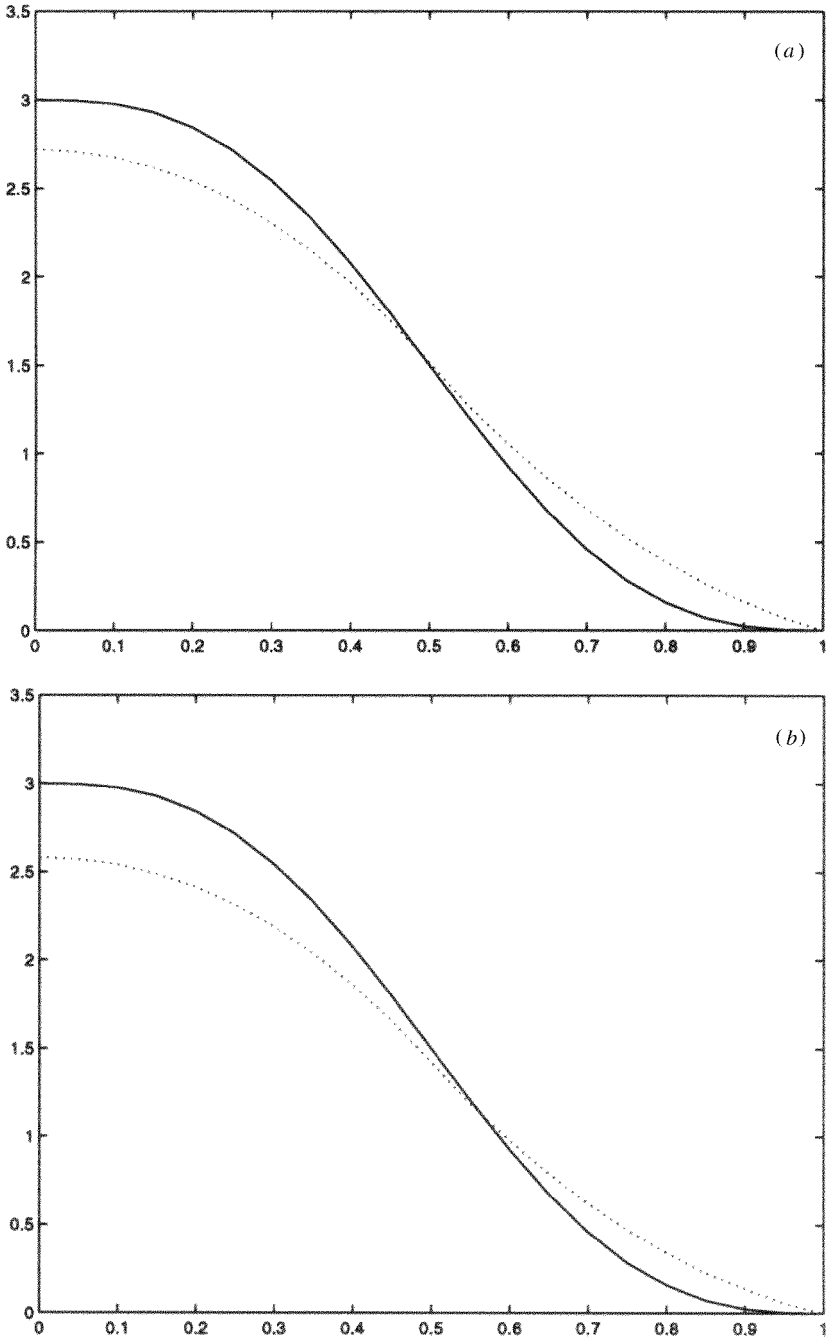

Figure 2. Exact solution (full) and iterate at $k = k_*$ (dotted) versus u, $\delta = 0.025$, $f = 50$. (*a*) corresponds to an experiment with temperature values in the whole range [0, 1], (*b*) to an experiment with temperature not decreasing below 0.2 (see figure 1).

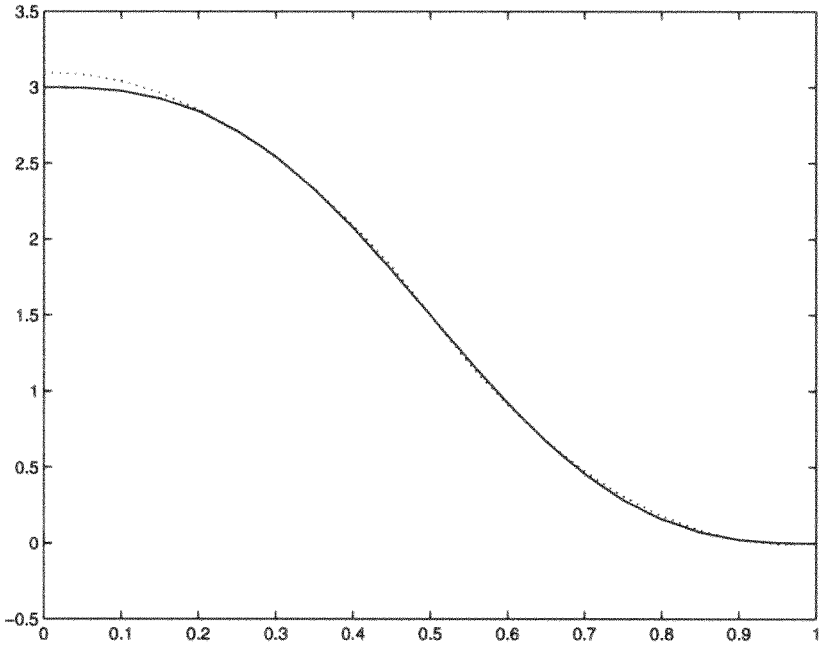

Figure 3. Exact solution (full) and iterate at $k = k_*$ (dotted) versus u, $\delta = 0.01$, $f = 30$, $a = 1$.

3. Numerical results

For the sake of simplicity we perform first numerical experiments for constant a ($a = 1$, $a = 0.5$ and $a = 0.05$), which has technical rather than conceptual reasons. The behaviour is similar to the relevant case of a bounded away from 0 uniformly, i.e. if a positive real number a_0 exists, such that

$$a(u) \geqslant a_0 \qquad \forall u \in ([u_1, u_2]).$$

The 'exact data' are computed by solving the direct problem with a very fine discretization and then perturbed by adding a high-frequency data error of L^2-norm δ (with frequencies $f = 30, 50, 100$ and an error level between 0.5 and 5%). The perturbed data u_B^δ and v_*^δ are obtained by sampling using a different discretization. We perform two kinds of experiments, the first with temperature values in the whole range, i.e. $[u_1, u_2] = [0, 1]$, and the second with a relevant temperature range $[u_1, u_2] = [0.2, 1]$ (see figure 1). The 'exact' rate b we use is similar to the (rare) measurements made for nucleation rates (cf [6]).

The results for the case $[u_1, u_2] = [0.2, 1]$ are illustrated at the right-hand side of figure 2. In this case, a source condition cannot be satisfied, because our starting iterate does not satisfy the boundary condition (2.36), so the convergence may be arbitrarily slow. Nevertheless, the approximation in the accessible temperature range is acceptable; a comparison with the left part of figure 2, which illustrates the corresponding result for $[u_1, u_2] = [0, 1]$, shows that the difference between exact and approximate solutions is larger only in the missing temperature range $[0, 0.2)$. For all further tests and the numerical confirmation of the predictions made by the theory we consider the case $[u_1, u_2] = [0, 1]$, with a starting iterate satisfying the source condition (2.30) or at least the necessary condition (2.35).

For the discrepancy principle, we use a relatively small parameter τ, here $\tau = 2.1$. As in figure 4, the residual in our computations usually did not decrease significantly below the value

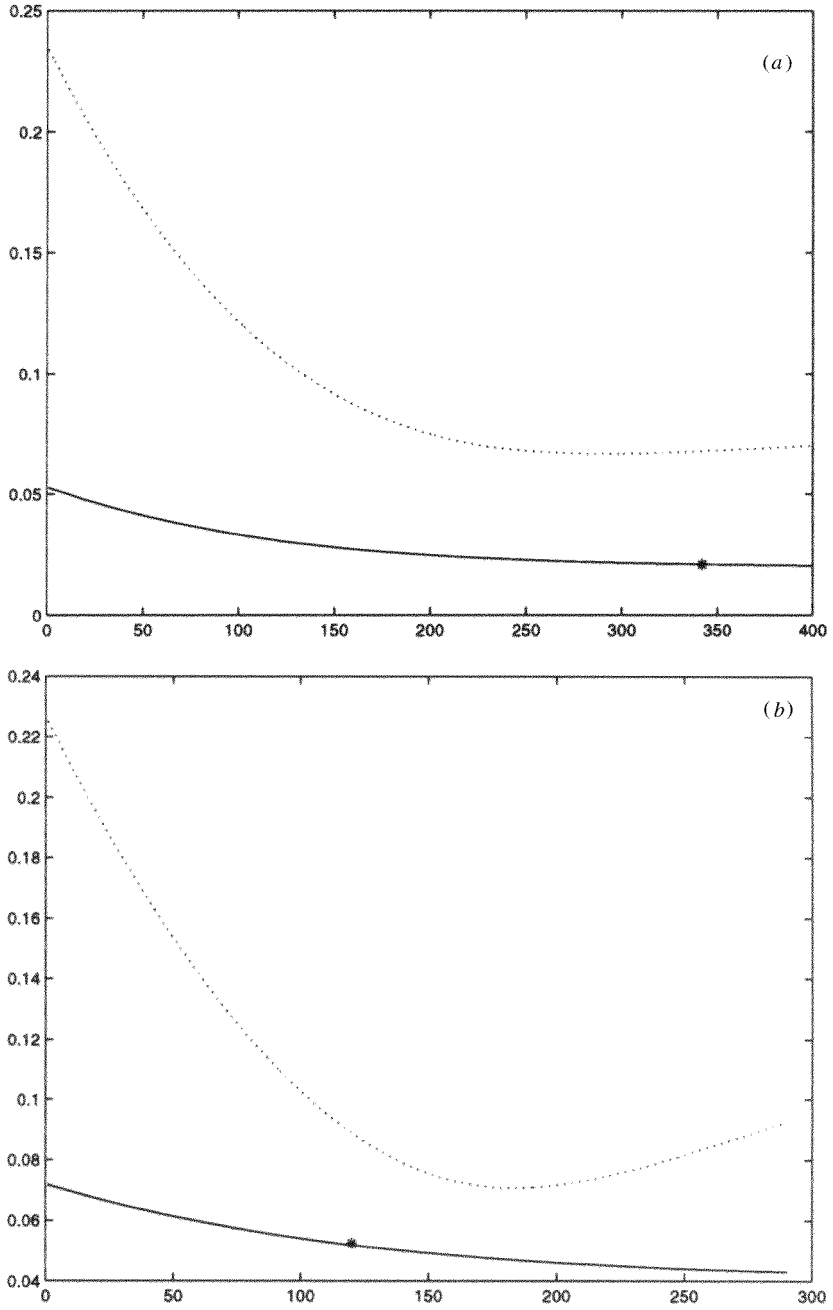

Figure 4. Residual (full) and error $\|b^k - b\|$ versus iteration number for $a = 0.5$, $\delta = 0.01$ (*a*) and $\delta = 0.025$ (*b*), $f = 50$. The stopping index is marked by $*$.

2δ. The numerical experiments show that the discrepancy principle (2.25) with $\tau = 2.1$ yields good results, a comparison of the iterates around the stopping index shows that the choice of the stopping index is not too critical for $a = 0.5$ and $a = 1$. This is due to the slow convergence

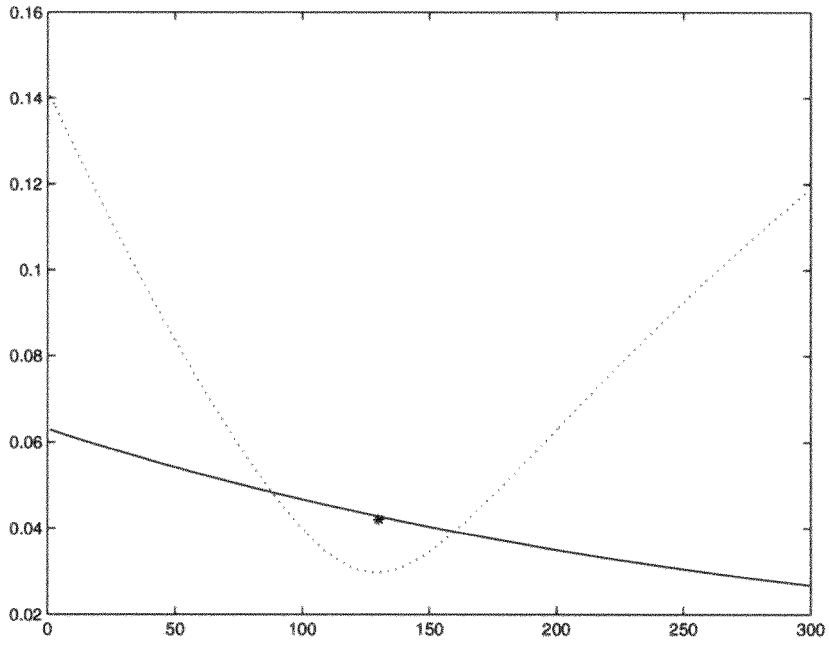

Figure 5. Residual (full) and error $\|b^k - b\|$ versus iteration number for $a = 0.05$, $\delta = 0.02$, $f = 50$. The stopping index is marked by $*$.

speed of Landweber iteration and would probably be more critical for Newton-type methods. The choice of the stopping index is more critical in the case of a 'small' parameter a, which may occur in practice (cf [6]), in our computations $a = 0.05$. This demonstrates again the ill-posedness of the problem, as figure 5 shows, only the first few iterates are close to the solution, a choice of τ less than 2 would lead to a high error between approximate and exact solution. Nevertheless, the solution determined using the discrepancy principle is always close to the iterate with minimal error in the H^1-norm, which demonstrates again the importance of a good stopping rule.

The damping factor ω clearly depends on the size of a, for our choice of the parameter a and b the method works without damping, although a line search strategy seems recommendable for accelerating the convergence speed. In the case of $a = 0.05$ we found that the parameter choice $\omega > 1$ leads to higher convergence speed. Nevertheless, one should not use a high ω if the residual is close to $\tau\delta$, because the faster convergence will also decrease the number of iterates close to the solution, which makes the choice of a good stopping index even more difficult.

As $\delta \to 0$, one observes that the approximations b^{k_*} converge to the exact solution, which is also illustrated by the results for a noise level of 2.5% at the left-hand side of figure 2 and for a noise level of 1% in figure 3. The residual and the difference between exact and approximate solutions develop as predicted by the theory. In particular one observes that the difference to the exact solution tends to zero as $\sqrt{\delta}$ and the stopping index behaves as $\frac{1}{\delta}$ (see figure 6), i.e. (2.31) and (2.32) are confirmed numerically.

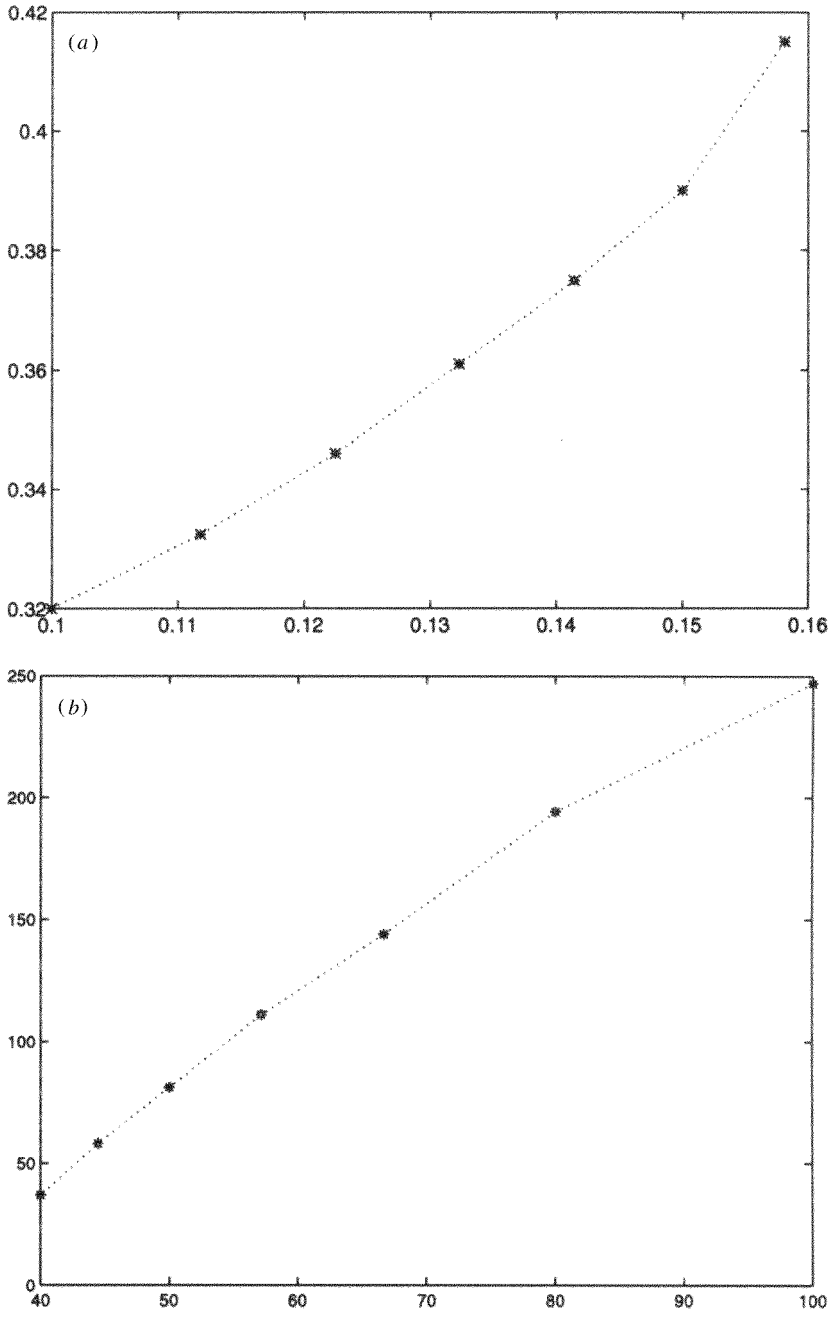

Figure 6. Error $\|b^{k_*} - b\|$ versus $\sqrt{\delta}$ and stopping index k_* versus $\frac{1}{\delta}$.

4. Extensions

As with the identification of nucleation rates, growth rates can be determined using Landweber iteration. The nonlinear operator that maps the parameter onto the data is given by

$$G : \mathcal{D}(G) \subset H^1([u_1, u_2]) \to L^2(I) \times L^2(\Omega) \tag{4.1}$$

$$a \mapsto (u|_{x=0}, v|_{t=t_*}). \tag{4.2}$$

If neither a nor b are known, it is necessary to identify them both simultaneously. The development of algorithms for this problem might be an important subject for future investigation.

The method presented above enables the numerical identification of the kinetic parameter in the one-dimensional model (1.1)–(1.7). We emphasize again that some of our analysis was formal, some technical mathematical details are still missing.

Another important extension is the case of two- and three-dimensional crystallization, where no partial differential equation formulation for the evolution of the degree of crystallinity seems to exist yet (cf [6]).

Acknowledgments

This work has been partly supported by the Christian Doppler Forschungsgesellschaft, Vienna, by the Austrian Fonds zur Förderung der wissenschaftliche Forschung, projects P 10866-TEC and SFB F 013, by ASI, contract ARS-96-121, and by the EU under the TMR-Network *Differential Equations in Industry and Commerce*.

References

[1] Andreucci D, Bianchini M and Pasquali A 1995 Identification of parameters in polymer crystallization *Appl. Numer. Math.* **17** 191–211

[2] Andreucci D, Fasano A, Primicerio M and Ricci R 1996 Mathematical problems in polymer crystallization *Surv. Math. Ind.* **6** 7–20

[3] Banks H T and Kunisch K 1989 *Estimation Techniques for Distributed Parameter Systems* (Basel: Birkhäuser)

[4] Binder A, Hanke M and Scherzer O 1996 On the Landweber iteration for nonlinear ill-posed problems *J. Inverse Ill-Posed Problems* **4** 381–9

[5] Blaschke B, Neubauer A and Scherzer O 1997 On convergence rates for the iteratively regularized Gauss–Newton method *IMA J. Numer. Anal.* **17** 421–36

[6] Burger M, Capasso V and Eder G 1998 Modelling non-isothermal crystallization of polymers *Working Paper* Industrial Mathematics Institute, University of Linz

[7] Cannon J R, DuChateau P and Steube K 1990 Unknown ingredients inverse problems and trace-type functional differential equations *Inverse Problems in Partial Differential Equations* (Philadelphia, PA: SIAM) pp 185–200

[8] Capasso V, DeGiosa M and Mininni R 1995 Asymptotic properties of the maximum likelihood estimators of parameters of a spatial counting process modelling crystallization of polymers *Stoch. Anal. Appl.* **13** 279–94

[9] Colton D, Ewing R and Rundell W (ed) 1990 *Inverse Problems in Partial Differential Equations* (Philadelphia, PA: SIAM)

[10] Deuflhard P, Engl H and Scherzer O 1998 A convergence analysis of iterative methods for the solution of nonlinear ill-posed problems under affinely invariant conditions *Inverse Problems* **14** 1081–106

[11] Eder G 1997 Fundamentals of structure formation in crystallizing polymers *Macromolecular Design of Polymeric Materials* ed K Hatada *et al* (New York: Dekker) 761–82

[12] Eder G, Janeschitz-Kriegl H and Liedauer S 1990 Crystallization processes in quiescent and moving polymer melts under heat transfer conditions *Prog. Polymer Sci.* **15** 629

[13] Engl H, Hanke M and Neubauer A 1996 *Regularization of Inverse Problems* (Dordrecht: Kluwer)

[14] Engl H, Kunisch K and Neubauer A 1989 Convergence rates for Tikhonov regularization of nonlinear ill-posed problems *Inverse Problems* **5** 523–40

[15] Engl H, Scherzer O, Yamamoto M 1994 Uniqueness and stable determination of forcing terms in linear partial differential equations with overspecified boundary data *Inverse Problems* **10** 1253–76

[16] Hanke M 1997 A regularization Levenberg–Marquardt scheme, with applications to inverse groundwater filtration problems *Inverse Problems* **13** 79–95

[17] Hanke M, Hettlich F and Scherzer O 1995 The Landweber iteration for an inverse scattering problem *Proc. 1995 Design Engineering Technical Conf. (Part C, Vibration, Control, Analysis and Identification)* vol 3, ed K W Wang (New York: ASME) pp 909–15

[18] Hanke M, Neubauer A and Scherzer O 1995 A convergence analysis of the Landweber iteration for nonlinear ill-posed problems *Numer. Math.* **72** 21–37
[19] Isakov V 1990 *Inverse Source Problems* (Providence, RI: American Mathematical Society)
[20] Liedauer S, Eder G, Janeschitz-Kriegl H, Jerschow P, Geymayer W and Ingolic E 1993 On the kinetics of shear induced crystallization in polypropylene *Int. Polymer Process.* **8** 236–44
[21] Pilant M and Rundell W 1990 Undetermined coefficient problems for quasilinear parabolic equations *Inverse Problems in Partial Differential Equations* pp 165–85 (Philadelphia, PA: SIAM)
[22] Ratajski E, Janeschitz-Kriegl H 1996 How to determine high growth speeds in polymer crystallization *Colloid Polymer Sci.* **274** 938–51
[23] Scherzer O 1995 Convergence criteria of iterative methods based on Landweber iteration for solving nonlinear problems *J. Math. Anal. Appl.* **194** 911–33